Self-Assessment as a New Challenge in Iranian Context:

Its Strengths and Weaknesses

Guiti Afrooz Javadi

University of Tehran

Introduction

"Teaching involves assessment."(Rea Dickenson, 2004)

In every day classroom situation almost all the interactions between teachers and learners can be considered as a form of assessment but there is a tendency to prioritize procedural and formal tests over assessment with a specific reference to teaching and learning and alternative forms of assessment. Classroom-based assessment has been looked down and not taken seriously by many educational authorities (Alderson, 2001). In educational settings there is an emphasis on measuring knowledge and academic achievement mostly through formal tests and examinations; therefore, teacher-made tests are the tests which are regularly practiced and the alternative assessments especially peer assessment and self-assessment are not very frequently applied to our educational settings.

Despite their professed eagerness to learn a language, EFL students are mostly prone to discontinue the endeavor shortly after their formal education is over and the evidence is the true life story of many of them when one encounters them some years after their graduation. It seems that in order to foster a life-long ability for pursing language learning in students, even long after they have finished their formal education, there should be a reorientation of the academic philosophy of language teaching and learning processes leaning towards independent learning and assessing.

Self-assessment seems to be a key component to be used in academic programs along with self directing learning. It can accompany other forms of assessment to provide teachers and decision makers with a more comprehensive picture of the learners' language abilities and shortcomings.

In teaching language to adults, it is important to accurately assess and then place learners in the levels and learning environments most beneficial to them. In order to achieve this, the reliability and validity of the tool for conducting such an assessment as well as ease in administration and interpretation should be considered (Bachman, 2006)

In addition to usual standard, paper and pencil tests, the learners' own assessment of

their abilities and weaknesses as well as their perceived achievements can reveal learners' confidence in their language ability. It can also assist in making placement tests, raise learners'

awareness of their strengths and weaknesses, and consequently foster autonomy in learning and assessing (Oscarsson, 1989).

Self-assessment is a meaningful and authentic assessment that gives the teachers a chance to assess their students on the basis of a wider range of evidence than the traditional tests. The evidence is provided by both teacher's assessment and learners' self-assessment. Self-assessment can also help in finding out whether the students are becoming competent and purposeful language users or not. As a result, language programs become more responsive to the individual learners' needs and learning strategies. (Council of Europe, 2001).

The results of self-assessment by adult learners have shown to highly correlate with assessments made by a variety of external means (Council of Europe, 2001). It is now recognized that learners do provide meaningful input into the assessment of their own performance, and that this assessment can be valid. Research reveals an emerging pattern of consistent, overall high correlations between self-assessment results and ratings based on a variety of external criteria (Blanche, 1988,; Oscarsson, 1984, 1998; Coombe, 1992, cited in Coombe.2001).

In fact, the question is: do the results of self-assessment done in non-Iranian contexts apply to Iranian contexts as well? And if yes, how should it be implemented? If self-assessment is to play a part in language programs, ways of assessing its value must be explored together with strategies for its effective implementations. Such an exploration will raise questions in regard to the assumptions and traditions underpinning other forms of assessments currently in use. One of the most important aspects of such an investigation will be the aspect of the attitudes of the people involved, namely, learners, teachers and policy makers.

This study aims to investigate the underpinning assumptions and attitudes of people involved in language learning and teaching, namely, learners, teachers and policy makers towards a specific form of alternative assessment, that is, self-assessment in an Iranian context. It also looks for a relationship or correlation between the results of the traditional tests and self- assessments from the validity, reliability and generalizability points of view.

Review of Literature

> It is suggested by Bachman (1989) that a reform and change in language education was brought about by a shift of focus from the behaviorist psychology and structural linguistics toward cognitive psychology and more meaning –based views of language. This change necessitates a reform in the type of teaching and learning approaches to be implemented as well as the type of assessment and evaluation of the learners' language achievement and proficiency   (Bachman, 1989, p.14).

Alternative assessment is the name chosen for the type of tests fit for such an intention. Alternative assessment has become an umbrella term for anything other than standardized, multiple choice questions. Examples include short answer response and extended response, individual or group performance assessment. Types of alternative assessments are tutor assessment, peer assessment and self-assessment.

A major topic in applied linguistics is the assessment of L2 proficiency and understanding and investigating developmental sequences of how different levels of proficiency are acquired. ((Thomas, 1994). Despite the shortcomings one observes in self-assessments. (Heron, 1988; Bachman and Palmer, 1989; Oscarsson, 1997; Brindley, 2001; Chapelle and Brindley, 2002), it is a widely adopted measure of the level of language proficiency in research studies that may contribute to this function.

As Bachman (2006) points out, research supports the claim that self-assessments, in low stakes situations, such as in applied linguistics can be both reliable and valid. But he immediately adds that like any other type of assessment, it must be based on a clear definition of the construct to be assessed, and then developed following sound test development practice. Bachman also recommends that it is much better to use multiple outcome measures rather than relying on only one.

Once the decision is made that alternative assessments are needed, one should be able to understand the merits and promises of using alternative assessments as well as the disadvantages and challenges inherent in such a practice especially when high-stake decisions (such as student retention, university entrance, promotion, graduation, and assignment to particular instructional groups) which have profound consequences for the students are concerned.

The change and reform in assessment system will become a tool for eliminating educational inequity only if it is accompanied with three other equally important factors, the ongoing professional development in curriculum and instruction, improved pedagogy, and quality assessment.

Liu and Littlewood (1997, p.79) point out that one of the reasons that educators are not so enthusiastic about using standardized multiple –choice tests, secured tests and other norm-referenced tests assessments is that all interpretations and inferences are made solely on the basis of these tests and this creates inequity. Sometimes a single test is the

only criterion for placing the student in a certain class or level, requiring the student to repeat the course, or even denying him the graduation or entrance to college or university.

At all levels educators are turning to alternative, performance- based assessments that are backed by criterion-referenced standards. Such assessments can help educators gain a deeper understanding of student learning, and communicate evidence of such learning to parents, employers and community at large. (Oscarsson, 1984, p.133).

The new alternative assessments and standards have been proposed as a cure to the problems observed in the educational systems including the gap in performance between the students of different ethnic, socioeconomic, and language backgrounds. This advantage of alternative assessment is not of great importance in an Iranian context as the importance of the type of assessment that would help the language learner to become an independent and autonomous language learner to pursue his learning after his official language courses are finished and he is on his own to further his language knowledge and ability (LeBlanche, 1985).

The goal of language education in my belief should be to enable and empower learners by providing enough training in self evaluation to sasses their own learning even when there is no teacher to assess them and tell them if they have learned what they were studying. Findings from cognitive psychology (cited in Ross,1998;Thompson,1996;Van Lier,1996; Thomas,2006) on the nature of meaningful, engaged learning support the use of alternative assessment that are tied to curriculum and instruction and that emphasize higher-order thinking skills and authentic tasks. Alternative assessments often require students to solve complex, real- life problems.

Some educators (Kumaradievelu, 1994, 2003; Oscarsson, 1984, 1997) believe that alternative assessments motivate students to show their best performance-performance that might have been masked in the past by standardized fixed –response tests and by uninteresting content. But it is a mistake if educators assume that by an exchange one high-stakes test for the other the equity will be achieved or the performance gap between learners will be eliminated (Gipps and Murphy, 1994).

Hammond (1994) believes that if new forms of assessment are to support real and lasting reforms and to close the achievement gap between students, they must be developed carefully and used for different purposes than the norm –referenced tests that have preceded .These purposes must be made explicit before the assessment system is built. She believes that this change mostly depends on the ways alternative assessments are used.

Alternative assessments are powerful tools for understanding student performance, particularly in areas that require critical thinking and complex problem solving but until the high expectations for success, sufficient opportunity to learn, and challenging instruction are the standard educational fare for all learners, some evidence (Elliot, 1993; Lemahieu, Eresh &Wallace, 1992) suggest that alternative assessment may reveal even greater achievement gaps than standardized assessments.

On of the most liberating and exciting things about the current interest in assessment is the recognition that that numerous assessment tools are available to schools ranging from standard fixed-response tests to alternatives such as performance assessment, exhibitions, portfolios, and observation scales .Each type of assessment brings with itself different strengths, and weaknesses to the problem of fair and equitable assessment. (McNamara, 2001, p.334)

Recognizing the complexity of understanding performance or success for individuals, it is virtually impossible that any single tool will do the job of fairly assessing student performance. Instead an assessment system made up of multiple assessments (including norm-referenced or criterion-referenced assessments, alternative assessments, self-assessments and classroom assessments) can produce "comprehensive, credible, and dependable information upon which important decisions can be made about students, schools, districts, or country."(Kolesch, Estrin, and Farr, 1995; Bachman, 2006).

Multiple assessment indicators are especially important for assessing the performance of ethnic-minority and language-minority students. The real challenge comes in selecting or developing a combination of assessments that work together as part of a comprehensive assessment system to assess all students equitably within the school community (Gipps, 1994.p.17).

The first and most critical step in assessing with equity is determining the purposes for assessing and clarifying whether those purposes are low stakes or high stakes (Winking & Bond, 1995). In many cases there may be multiple purposes-some low stakes and some high stakes – for assessing student performance. Based on the discussions above and havi**n**g a schematic view of the Iranian system of language teaching, it seems logical to assume that there is an urgent need for a change is in the field of second and foreign language teaching and testing in Iran. Over the past forty years a move away from the tenets of behaviorist psychology and structural linguistics toward cognitive psychology and more meaning- based views of language is observed.(Jackobs,2001,pp.4-8)

1. **A greater focus of attention on :** learners rather than teachers ,on learning process rather than on the products that learners produce, on the social nature of learning rather than on students as separate, decontextualized individuals, on diversity among learners and viewing these differences not as impediments to learning but as resources to be recognized, catered to and appreciated (study of individual differences), on the views of those who are internal to the classroom rather than valuing the views of outsiders who come to observe the class. This shift has led to qualitative research –with its valuing of the subjective and affective of the participants' insider views and of the uniqueness of each context.

2. **Emphasis on the context**: This emphasis led to the idea of connecting the school with the world beyond as a means of promoting holistic learning.

3. **Student Awareness**: Helping learners to understand the purpose of learning and develop their own purposes.

4**. A whole-to-part orientation:** A whole to part orientation instead of a part-to-whole approach which involves such approaches as beginning with meaningful whole texts and then helping

students understand the various features that enable the text to function, e.g. the choice of words and the organizational structure of the text.

5. Focus on the meaning: An emphasis on the importance of meaning rather than drills and other forms of rote learning.

6. Lifelong learning: A view of learning as a lifelong process rather than something done to prepare for an examination. (Jacobs, 2001)

Moreover, Jacob (2001, pp.9-10) gives six features regarding the formation of an approach to learning and assessment. They are:

1. Students will be involved not only in planning but also in assessing their own learning
2. Classroom assessment is understood, used and communicated among the teachers and educators with the ultimate result of supporting learning.
3. Assessment becomes an integral part of instructional planning.
4. Teachers and students learn and practice self-reflection
5. Teacher training programs are geared towards including classroom assessment practices to prepare teachers to undertake such activities to support learning for their pupils.
6. Parents also get to know and understand classroom assessment and communication practices that support learning. (Jacobs,2001)

Altogether the important changes mentioned above have affected potentially the language teaching practice and learning concepts. If these changes are implemented in an integrated manner they have the potential of converting our practice dramatically. These changes are: Learner autonomy, cooperative learning, focus on learning, thinking skills, curricular integration, diversity, alternative assessment, and teachers as co-learners.

The focus of my thesis is going to be specifically on an example of alternative assessment, that is, self-assessment. It is suggested that if self-assessment training is given to the learners and teacher trainees it can foster autonomy in learners.

Learner autonomy is related to Vygotsky's (1978) concept of self-regulation and Csikszementmihly's (1990) work on flow. Autonomous learners have some choice as to the "what and how" of the curriculum; at the same time they feel responsible for their own learning and for the learning of those with whom they interact. Learner autonomy involves Learners being aware of their own ways of learning, so as to utilize their strengths and work on their weaknesses (Van Lier, 1996).

The concept of learner autonomy fits with the overall paradigm shift because it emphasizes the role of the learner rather than the role of the teacher. It focuses on the process rather than the product and encourages the students to develop their own purposes for learning as a lifelong

process. The concept of autonomy has certain implications for second language teaching: For example the use of small groups—including pairs—is one means of enhancing learner autonomy (Harris& Nyau, 1990).

Intrinsic motivation plays a central role in learner autonomy. The teacher no longer shoulders the entire burden of the running the classroom. A form of democratization takes place with students taking on more rights and responsibilities for their own learning (Littlewood, 1999). Another implication for Second language education flowing from the concept of learner autonomy is extensive reading. Here, students choose reading material that matches their own interests and proficiency level. If a student begins a book or a magazine and it does not seem the right one for him or her, he/she can switch to another. The hope is that extensive reading will aid students in developing an appreciation for the enjoyment and knowledge to be gained via reading in their second language (as with their first), thus encouraging them to make reading a lifelong habit.

Self- assessment is yet another way that provides learners with autonomy in their learning. (Lee, 1998; Rothschild& Klingenberg, 1990).The idea is for the learners to develop their own internal criteria for the quality of their work, rather than being dependent on external evaluation, often by the teacher, as the sole judge of their strengths and weaknesses. Developing these internal criteria enables learners to make informed decisions about how to move their learning forward.

With self-assessment, no longer do students have to wait for the teacher to tell them how well they are doing and what they need to do next. Although the teacher remains generally the more knowledgeable and experienced person in the classroom, but the goal is for students to move toward and perhaps even beyond, the teacher's level of competence. Placing value on learners' knowledge helps them feel more capable of playing a larger role in their own learning. (Oscarsson, 1984; Blanche, 1988).

Another paradigm shift relates to taking diversity of the learners into account, usually manifested in their learning styles and strategies. Among the strategies that learners need to acquire and use are those that involve going beyond the information given and utilizing and building higher –order thinking skills (Paul, 1995)). Today, thinking skills are seen as an essential part of education, because information is easily obtained, so the essential task is now to use that information wisely.

Connecting education to the wider world in order to improve that world means that students along with their teachers need to analyze existing situations, synthesize new ideas and evaluate the proposed alternatives (Freire, 1970).The concept of thinking skills flows from the current paradigm shift in a few senses. First, thinking is a process and the emphasis lies in the quality of that process rather than solely on the quality of the product resulting from that process (Li, 1998, Lee, 2001).

Additionally, many valid routes may exist to ward thinking about a particular situation. Another connection between thinking skills and current paradigm is the attempt to connect the school with the world beyond. This attempt promotes the idea that learning is not a collection of lower-order

facts to be remembered and then regurgitated on exams, but that we learn in school in order to apply our knowledge toward making a better world(Richards and Hart,1994).

Many attempts are being made to integrate thinking across curriculum. Group activities provide a useful venue for second language students to move away from sole reliance on forms of assessment that need higher –order thinking skills with questions that have more than one possible correct answer. Also projects and other complex tasks are being used for assessment purposes.

New alternative assessment instruments are being developed to complement or replace traditional instruments that use multiple choice, true-false and fill- in- the- blank items (Goodman, Goodman, &Hood, 1989). Further attempts are being made to develop assessment instruments that mirror real- life conditions and involve thinking skills. These alternative assessment instruments are often more time- consuming and costly, as well as less reliable in terms of consistency of scoring. Nevertheless, they are gaining prominence due to dissatisfaction with traditional modes of assessment, which are faulted for not capturing vital information about student's competence in their second language (Gipps and Murphy, 1994).

Alternative assessment has developed on many different fronts in second language education. One of these involves the teaching of writing. In the process of approach to writing, students go through multiple drafts as they develop a piece of writing. Rather than only evaluating the final draft, teachers now look at earlier drafts as well to gain a better understanding of the process students went through as they worked toward their final draft.

Portfolios offer a complementary means of looking at students writing processes ((Fusco, Quinn, and Hauck, 1994).With portfolio assessment, students keep the writing they have done over the course of a term or more, including early drafts. Then, they analyze their writing to understand the progress they have made. Next, they select from among their pieces of to compile a collection that demonstrates the path of their writing journey and prepare an introduction to the portfolio in which they present their findings.

Another alternative form of assessment in second language education involves peer assessment (Chen, W., & Warren, 1996) .This form of assessment is intended to enhance, not replace self- and teacher assessment. Peer feed back is now common in writing classes. By critiquing the writing of fellow students, learners better understand and internalize criteria for successful writing.

Under the old paradigm, teaching is seen as a skill that can be learned in discrete items from Lesson planning to how to ask questions. When these skills are learned, the teacher is qualified to teach. In second language teaching this approach to education is seen as "training". (Freeman, D. (1996). However, the current paradigm sees the teacher and the learning process as social processes where the students are active constructors of knowledge with their teachers. The teacher is more of a facilitator and fellow learner alongside the students.

In the previous paradigm, second language teachers' opinions and experience were often not included. Instead, the experts in the universities did the research and administers did the assessment. Their pronouncements were then handed down to practitioners. In the current paradigm, the notions of qualitative, ethnographic research by and with the teachers and self and peer assessment of teachers has unfolded. ((Fanselow, 1987).

Cooperative learning (CL) connects with learner autonomy because group activities help second language learner s become less dependent on teachers. Alternative assessment is fostered in several ways by the use of CL. For instance, CL provides scope for peer assessment and an emphasis on the development of collaborative skills calls for different methods to assess these skills.

The question is: Has this paradigm shift affected the second language classrooms? It seems that the effect of such a shift is only being slightly felt in our country. There seems to be a great deal of variation between institutions, and even classrooms within the same institution. Thus, the so called rapid, revolutionary, far-reaching paradigm shifts in the physical sciences turns into a gradual, evolutionary and piecemeal one in second language teaching. The reasons for such a difference can be described as:

➢ Changing beliefs and behaviors takes time in education and elsewhere.

➢ Lack of change may be the result of difficulty of translating theory into practical application. That is new ideas need a great deal of work by practicing teachers to translate into their every day teaching routines.

➢ Another possible explanation for the lack of implementation of this paradigm shift stems from the fact that it has often been presented in a piecemeal fashion, rather than as a whole. ((Fullan, Bennett. &Roleheiser-Bennett, 1990)

The point is that many of the changes that we hear about in education in general and second language education in particular are all part of the one overall paradigm shift and trying to separate and perceive them as unrelated and in a piecemeal fashion will lessen our chances of success as well as understanding the changes themselves and applying them in practice. The innovations fit together, like the pieces of in a pattern cut to make a jigsaw puzzle. Each piece supports the others. Therefore, in order to increase the chances of understanding the changes in our profession, educators must take a big picture approach and implement change in a holistic way (Jacobs, 2001).

Unfortunately, in many cases while teaching methodology has become more communicative, testing remains with the traditional paradigm. The old paradigm consists of discrete items, lower order thinking and a focus on form rather than on meaning. This creates a backwash effect that tends to pull teaching back toward the traditional paradigm, even when teachers and others are striving to go toward the new paradigm.

To put it in a nutshell, it is a well known fact that implementing change is difficult. All the changes that have taken place in our field challenge us to continue learning about our profession and to share what we learn with others, including our colleague, so that we can continue to help our field develop. It is the hope of the writer of this proposal that the result of the research can shed a flicker of the light on the crucial issue of fair assessment.

References

Bachman, L.F. (2000). Modern language testing at the turn of the century: assuring that what we count counts. *Language Testing*, 17, 1-42.

Bachman, L. F. and Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing*, 6, 14-25

Boud, D. (1989). The role of self-assessment in student grading. *Assessment and evaluation in higher education*, 14, 20-30. (In electronic format at http://www.Finch Park.com/arts/self.html/, retrieved on 8/11/2006

Brown, H. D. (1994*). Principles of language learning and teaching* (3rd.ed).Upper Saddle River, NJ: Prentice Hall Regents.

Brindley, G. (2001) Assessment in Reading, Carter, and Nunan (eds)*)* The Cambridge Guide   To Teaching English to Speakers of Other Languages. Cambridge: Cambridge University Press.

Bruna, K. R. (2002), Screening and placement of students into appropriate Spanish levels: Self-assessment, cognitive screening, and survey data, 2002 (pp.12-15).

Cheng, W.and Warren. (1996). Hong Kong students' attitudes toward peer assessment in English Language Courses, *Asian Journal of English Language Teaching*,6,61-75.

Csikszentmihalyi, M. (1990).Flow: The Psychology of optimal experience. New York: Harper and Row.

Coombe, Christine. (1994). Self-assessment in language testing: E (merging) Strategies for Second/Foreign Language Teaching. *TESOL Quarterly*, Vol.28, No.1, 1994, 39-40

Fanselow, J., F. (1988). Let's see Contrasting conversations about teaching. *TESOL Quarterly,* 22,113-130

Farrell, T.S.C. (1999). Teachers talking about teaching: Creating Conditions for Reflection. *Teaching English as Second or Foreign Language*, 4(2), (in electronic format at:http://www-writing.berkley.edu/TESL-EJ/ej14/a/.html/

Freeman, D. (1989). Teacher Training, development, and decision making: A model of teaching and related strategies for language teacher education, *TESOL Quarterly*, 23, 27-45.

Fullan, M.G., Bennett. &Roleheiser-Bennett. (1990).Linking classroom and school Improvement, *Educational leadership,* 47(8), 13-190).

Freeman, D. (1996). Teacher training, development, and decision making: A model of teaching and related strategies for language teacher education. *TESOL Quarterly*, 23, 27-45)).

Fusco, E., Quinn, M., Hauck, M. (1994). The portfolio assessment handbook. Austin, TX: Steck Vaughn.

Goodman, K., Goodman, Y., Hood, W. (1989). The whole language evaluation book. Portsmouth, N.H.: Heinemann)).

Harris, V., &Noyau, G. (1990). Collaborative learning: Taking the first steps. In Gathercole, I. (Ed.), Autonomy in language learning (pp.55-64).)) London.

Harris, M. (1977) Self assessment of language learning in formal settings. In *ELT Journal* Vol.51/1, pp.12-20.Oxford: Oxford University Press.

Heron, J. (1988). Assessment revisited. In D. Boud (ed.), *Developing student Autonomy in Learning*. New York, N.Y.: Kogan Page.

Jacobs, M.G. (2001). Paradigm Shift: Understanding and Implementing change in Second Language Education, TESL-EJ, Vol.5, no.1, retrieved from: http://www-writing.berkley.edu/TESL-EJ17/al.html

Jafarpour, A. (1991). Can naïve EFL learners estimate their own proficiency? Evaluation and Research in, 5,145-*57*.

Kumaravadivelu, B. (1994). The post method condition: (e) merging strategies for second/foreign language testing. *TESOL Quarterly*, 28, 27-48

Kumaravaievlu, B. (2003). Beyond Methods: Macro strategies for Language Teaching. New Haven, C.T.: Yale University Press.

Lee, I. (1988).Supporting greater autonomy in language learning.*ELT Journal*, 52,282-290.

Lee, I. (2001). Hong Kong students' attitudes toward peer assessment in English Language Courses. *Asian Journal of English Language Teaching*, 6, 61-75.

Li, D., 1998. "It's always more difficult than you plan and imagine": Teacher's perceived difficulties in introducing the communicative approaches in South Korea. *TESOL Quarterly*, 32 (4), 677-700

Liu, N.F., Littlewood, W., (1997). Why do many students appear reluctant to participate? in classroom learning discourse? *System* 25(3), 371-384.

Oscarsson, M. (1997). Self-Assessment of foreign and second language proficiency. In C.clapham and D. Corson (eds.) Encyclopedia of Language and Education, Vol. 7: *Language Testing and Assessment.*

Oscarsson, M. (1978) Approaches to Self-assessment in Foreign Language Learning. Council of Europe: Pergamon Press

Paul, R.W. (1997)). Critical thinking: How to prepare students for a rapidly changing world. Santa Rosa, CA: Foundation for critical thinking.

Ross, S. (1998). Self-assessment in language testing: a meta analysis and measures of experiential factors. *Language Testing*, 15, 1-20

Thompson and Chihirokinoshito., Self-assessment in Self-directed Learning: Issues of learner diversity.Pemberton. Richard et al. eds.*Taking control: Autonomy in language Learning*. Hong Kong, 1996 (PP.77-91)

Thomas, M. (1994). Assessment of L2 proficiency in Second Language Acquisition Research*, Language Learning,* 44,307-336

Thomas, M. (2006), ILT, *Language Teaching Journal*, Vol.2, No.2.2006

Van Lier, L. 1996, *Interaction in the language curriculum: Awareness, autonomy & authenticity*. London: Longman.

Vygotsky, L.S. (1978). *Mind in Society*. (ed. By M.Cole, V.John-Steiner, S.Scribner and E.Souberman). Cambridge, MA: Harvard University Press.